1986

Memory for Tune Titles after Organized or Unorganized Presentation

Andrea Halpern
Bucknell University, ahalpern@bucknell.edu

Memory for tune titles after organized or unorganized presentation

ANDREA R. HALPERN
Bucknell University

Two experiments investigated the structure of memory for titles of 54 familiar tunes. The titles were presented in the form of a hierarchy, with nodes labeled by genre (e.g., *Rock* or *Patriotic*). Four groups of subjects received logical or randomized titles, and logical or randomized labels. Goodness of label and title structure had equal and additive beneficial effects on recall with a 3-min exposure of the stimuli. With a 4-min exposure, good title structure became a larger contributor to good recall. Clustering analyses suggested that subjects were mentally representing the tune titles hierarchically, even when presentation was random.

This paper concerns the way people learn and remember titles of familiar tunes. It is obvious that most people have some access to memories for the titles, melodies, or lyrics of hundreds of tunes. This information can be very durable, lasting 50 or more years (Bartlett & Snelus, 1980). Many of us have even been annoyed at the durability of these traces, as in the common complaint "I can't get that tune out of my head!" One key to good memory for items is their organization into a coherent mental structure. This organization may be preexperimentally defined (e.g., Bousfield, 1953, for word categories; Thorndyke, 1977, for stories), or subjectively imposed (e.g., Mandler, 1967), but in either case, organization and memory are consistently linked. Analogous to previous work using ordinary verbal material as stimuli, the current studies employ verbal referents to tunes as stimuli. The hypothesis was that links between organization and memory could be found in this domain as well.

Evidence for such links was found in a previous study on memory for familiar songs (Halpern, 1984). In Experiment 1 of that study, subjects sorted titles of familiar tunes into piles according to musical and then nonmusical criteria of their own choice. The nonmusical ADDTREE (Sattath & Tversky, 1977) clustering solution captured a high proportion of variance in the data and produced a hierarchy wherein many of the primary (e.g., *Popular*) and secondary (e.g., *Rock*,

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 1986, Vol. 99, No. 1, pp. 57–70

Beatles) branching groups could be reasonably labeled. The musical solution produced less satisfactory results.

To confirm the psychological reality of the nonmusical similarity structure, Halpern (1984) in Experiment 2 asked a new group of subjects to verify whether a sounded tune and a displayed title referred to the same song. Few errors were made when mismatched songs came from distant points in the ADDTREE diagram; more than twice as many errors were committed when tunes were close together in the diagram. In addition, a surprise free-recall task showed more frequent clustering of items close together in the ADDTREE solution than of items a medium or far distance apart.

These results suggest that people organize their musical memory hierarchically. The current experiments further explore the reality of this proposed hierarchy by using a learning task. If a large number of song titles are presented to subjects in a form similar to that produced in the sorting solution, will subjects learn them more quickly than if titles are presented in a less organized format?

Bower and Clark-Meyers (1980) asked a similar question about script activities. Scripts refer to the organized bits of stereotyped information we possess about common situations (e.g., going to the dentist, Shank and Abelson, 1977). Bower, Black, and Turner (1979) found that people were comfortable describing and remembering these routine scenarios as a hierarchy of scenes each containing appropriate actions. For instance, "reading a magazine" might be contained in a "waiting room" scene.

In their learning study, Bower and Clark-Meyers (1980) presented 84 words to subjects for study and recall. The words were divided into groups, each containing the name of a script ("Concert") and nine script-related items or activities ("Usher," "Conductor") in their typical order of occurrence in the script. The scripts were organized into "Morning," "Noon," and "Evening" clusters, and the entire structure emerged from a node labeled "Calendar."

Bower and Clark-Meyers (1980) found that recall of all the words was excellent when they were presented in the hierarchical fashion just described. However, when the words were completely randomized, recall was quite poor. The authors supposed that emergence of a readily apparent theme enabled subjects to use their script-based knowledge to guide learning and memory search. Presented with exactly the same items in a random order, subjects resorted to inefficient, idiosyncratic strategies to accomplish the task.

The current experiments adapt and expand the Bower and Clark-Meyers procedure to the learning of the song title hierarchy already described. Besides tapping a very different kind of knowledge, these

studies also investigate which parts of a hierarchy are most important in learning: the organization of the category labels, or the organization of the titles themselves. Four different groups of subjects received normal or randomized labels, combined with normal or randomized titles. The relative effectiveness of good label or title arrangements can thus be examined. In addition, the actual pattern of recall of labels and titles in the protocols should enable us to characterize learning of the tunes as hierarchical (using the category labels to aid recall) or listlike.

Another aspect of the learning task to be examined is the clustering of items in recall. If level of recall and amount of clustering covary, evidence linking organization and memory for tune titles would be strengthened. In addition, if the materials are being remembered hierarchically, clustering should be stronger for items close together and weaker for items farther apart.

EXPERIMENT 1

METHOD

Subjects

Participants were 32 Bucknell University undergraduates, all of whom volunteered for the study. All were born and raised in the United States.

Materials

Four charts were prepared. Each was lettered with black ink on yellow posterboard, 2 × 6 ft (0.6 × 1.8 m) in size. Category labels were enclosed by boxes. Song titles were listed underneath category labels. Vertical or oblique lines connected all the song titles in a category, and all the subcategories to higher level categories, in a hierarchical fashion.

The four charts resulted from factorially combining two schemes of category labels (logical or illogical) and two arrangements of the song titles (organized or scrambled). The logical-organized version was a modification of the clustering solution for the nonmusical sort found in a previous study (Halpern, 1984). From that diagram, 54 song titles were used, as were 16 category labels that had been generated post hoc.

The illogical label structure was generated by randomly placing the 16 labels in the same slots as in the logical version. Likewise, the scrambled title structure resulted from completely randomizing the song titles. The logical-organized chart is shown in Figure 1, and the illogical-scrambled chart is shown in Figure 2. The two remaining charts used the logical labels with the scrambled titles or the illogical labels with the organized titles.

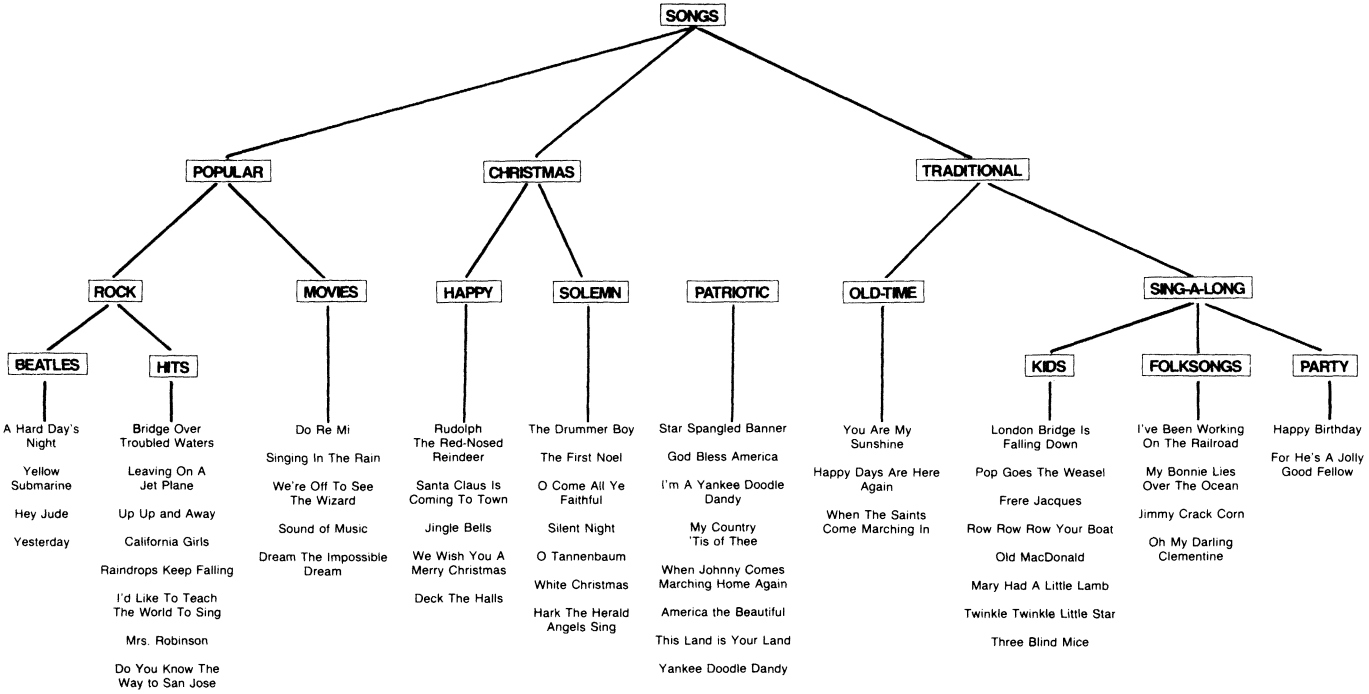

Figure 1. Logical (labels)-organized (titles) stimulus chart

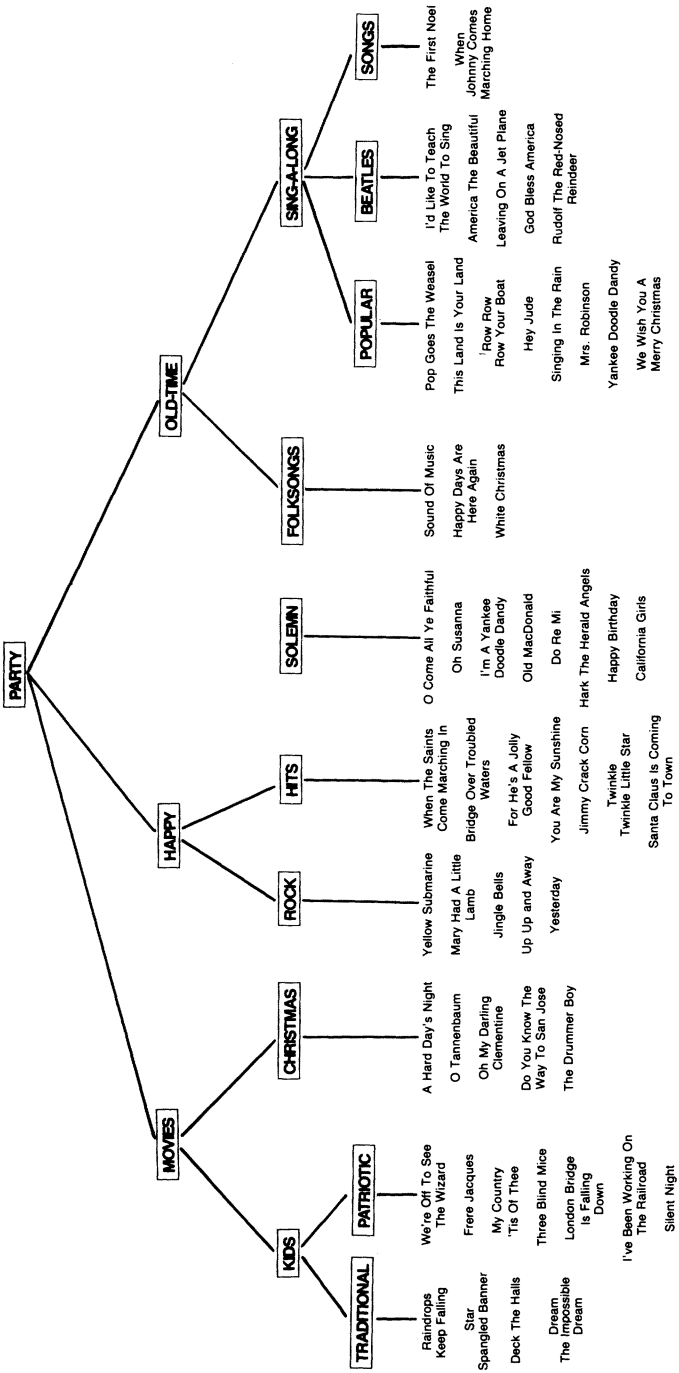

Figure 2. Illogical (labels)-scrambled (titles) stimulus chart

Procedure

Eight subjects saw each of the four charts. Four subjects were tested at a time. The covered chart was placed at eye level, approximately 5 ft (1.5 m) from the participants. After instructions to study the chart (when revealed) for a subsequent memory test, the chart was uncovered for 3 min. After it was covered again, subjects wrote down all the titles and labels they could remember in any order they wished, as long as recall was in a continuous list. They were given 6 min for recall. The study-test cycle was repeated twice. For each recall test, subjects were instructed to write down all the items they could remember, including ones they had recalled on a previous trial. Subjects were not allowed to consult previous recall attempts on a given trial.

RESULTS AND DISCUSSION

The primary measure of interest is the mean number of items recalled in each group for each trial. These data were analyzed via a three-way ANOVA, with labels (logical or illogical) and titles (organized or scrambled) as between-subjects factors, and trial (1, 2, or 3) as the within-subjects factor.

A strong learning effect was shown as an increase of recall over trials: of 70 items (16 labels and 54 titles), 17.5 (1.0 label and 16.5 titles), 27.2 (1.9 labels and 25.3 titles), and 36.2 (2.4 labels and 33.8 titles) were recalled on Trials 1, 2, and 3, respectively, $F(2, 56) = 248.8$, $p < .001$, for the main effect of trials. This factor did not interact with either of the other factors, so the remaining results will be discussed as collapsed over trials. In addition, analyses done separately for recall of labels and titles showed essentially the same patterns. Therefore, further results will be reported as recall of total items unless otherwise noted.

Logical labels engendered better recall than illogical ones, $F(1, 28) = 9.6$, $p < .01$, and recall of organized titles exceeded that of scrambled ones, $F(1, 28) = 14.6$, $p < .001$. These two factors did not interact, $F < 1$, as may be seen in the dark lines of Figure 3. The level of recall of charts with either logical labels or organized titles was equivalent; it was intermediate compared with the logical-organized or illogical-scrambled versions. As may also be noted in Figure 3, the size of the two main effects was about equal. That is, a logical arrangement of just 16 category headers aided recall as much as did a logical arrangement of the 54 song titles, relative to their respective unorganized versions. The illogical-organized chart, with many pieces of "good" information (title relationships), should perhaps have been predicted to produce a higher level of recall than the few pieces of

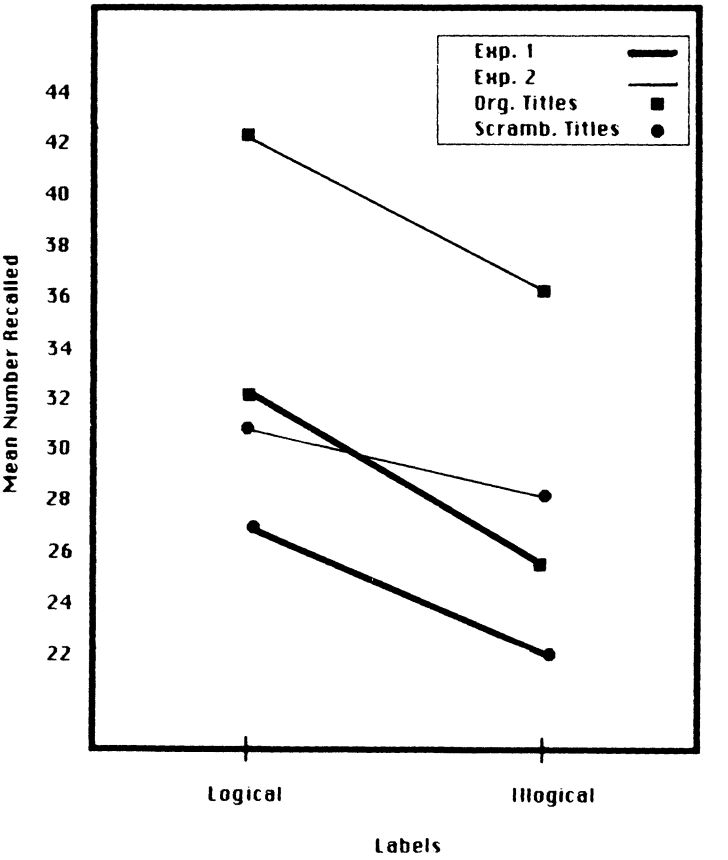

Figure 3. Mean number of items recalled for each chart in Experiments 1 and 2 (maximum = 70)

"good" information (label relationships) in the logical-scrambled condition. The similarity of recall levels for these two charts suggests that subjects were compensating for the irrationality of one component by using the other one.

Although recall level in this experiment was adequate, the next study increased stimulus exposure time so as to bring recall performance nearer to the levels found by Bower and Clark-Meyers (1980). This might also allow more opportunity for the learners of the more intelligible charts to exploit that organization in recall. In addition, data in Experiment 2 were analyzed more thoroughly for the way in which subjects were recalling each diagram (these data were unavail-

able for further analysis in Experiment 1). To this end, clustering patterns and use of category labels in recall were examined.

EXPERIMENT 2

METHOD

Subjects

An additional 32 Bucknell University students participated in this experiment. They met the same criteria for selection as those in Experiment 1.

Materials

The same stimuli as in the previous experiment were used.

Procedure

The procedure was the same as in Experiment 1, except the chart was exposed for 4 min on each trial, and subjects were tested in pairs.

RESULTS

Increasing the exposure time increased the level of recall, mostly for labels. Of 70 items (16 labels and 54 titles), an average of 22.5 (8.1 labels and 14.4 titles), 35.3 (9.8 labels and 25.5 titles), and 44.8 (11.3 labels and 33.5 titles) were recalled on Trials 1, 2, and 3, respectively. Recall on Trial 3 with the logical-organized chart averaged 77% (compared with 60% previously); performance on Trial 3 of the illogical-scrambled chart improved to 53% (from 43%).

As in Experiment 1, the amount of recall was analyzed by a three-way ANOVA. The learning effect referred to in the previous paragraph was confirmed, $F(2, 56) = 531.8$, $p < .01$, for the main effect of trials. Also similar to Experiment 1 were the main effects of titles, $F(1, 28) = 23.0$, $p < .001$, and labels, $F(1, 28) = 5.8$, $p < .05$, and the lack of a significant interaction between these factors, $F < 1$.

One difference between the two experiments was a Title × Trial interaction, $F(2, 56) = 5.5$, $p < .01$. The advantage in recall of organized over scrambled titles increased somewhat from Trial 1 (8.2 items) to Trial 2 (10.7 items) to Trial 3 (12.7 items). No other interactions were significant. Because the Title × Trial interaction does not significantly qualify the major results, recall performance was collapsed over trials and displayed as the thin lines in Figure 3.

A comparison of Experiments 1 and 2 reveals a number of similar results. The lack of interaction between titles and labels is particularly striking. Also, the advantage of logical over illogical labels is about

the same in both experiments. However, some differences are also apparent. Recall level is generally superior in Experiment 2, as previously noted. In addition, the advantage of organized over scrambled titles is considerably greater in Experiment 2. One final (but related) difference is that in Experiment 2, the illogical-organized chart produced better recall than the logical-scrambled chart (equivalent in Experiment 1). Overall, the longer stimulus exposure time in Experiment 2 particularly benefited recall for charts with the organized title scheme.

Category use

The next question explored was the difference, if any, in the nature of the recall strategies among the groups. One way in which the groups might differ is the extent to which labels aided recall of the titles. An efficient way to memorize and emit all these charts would be to mimic the organization of the logical-organized chart. That is, maximum use should be made of its hierarchical nature. Recall of a high-level heading, followed by a subheading, followed by its category members, another subheading, and so forth, would exploit the organizational and retrieval value of the label structure. It was predicted that recall protocols would most closely follow this pattern in the logical-organized condition, because of the ease of learning or emitting a scheme that not only makes sense but is also the actual stimulus presented. The illogical-scrambled chart was predicted to show the least amount of category-title recall, due to the difficulty of reorganizing both the label and title structure. Under these circumstances, category labels would be rendered ineffective as retrieval cues.

The two intermediate charts were predicted to show a little more category use than the completely nonrational (illogical-scrambled) chart. For the logical-scrambled chart, the rational labels could serve as retrieval cues for organizing the titles. Thus, at output, memory for the label "Movies" might trigger recall of appropriate titles. Because this organization would presumably occur mostly at output, its incidence would be less frequent than if the chart had been learned hierarchically. The illogical-organized chart might engender some categorical recall using a similar process. As a subject prepares to emit a group of related songs, the category label might come to mind and be inserted at the head of the group at output. However, the frequency of this strategy would be relatively low for the same reasons as noted for the previous chart.

For each subject's recall of Trial 1, the total number of labels was noted. From this was subtracted the number of times a label appeared without a category member appearing somewhere on the protocol in

Table 1. Proportion of category use for each chart in Experiment 2

	Labels	
Titles	Logical	Illogical
Organized	.38	.07
Scrambled	.06	.10

order to correct for opportunities to display appropriate label use. "Member" was defined as any title that was a direct descendant of the label in the logical-organized chart. For instance (referring to Figure 1), "Hey Jude," preceded by either *Songs, Popular, Rock,* or *Beatles* (its direct antecedents), would be counted as a positive instance. The proportion of times that a label was immediately followed by a category member was then computed and averaged (Table 1). (Subjects who recalled no labels on Trial 1, $n = 3$, were removed from the analysis to again correct for having opportunities to respond.)

As shown in Table 1, category use was by far the highest for the logical-organized chart. The other three charts did not differ much in the level of category/title recall they produced.

Clustering

The previous analysis considered the relationship of labels to titles in recall. The next analysis considered a more common description of recall patterns, clustering. The grouping of items in a free-recall protocol is commonly considered to reflect the degree to which items are grouped in memory. Considering only the title schemes in these charts (organized or scrambled), one may measure to what extent songs are recalled together according to the scheme of the organized diagram (hereafter referred to as "conceptual clustering"). If subjects can reorganize scrambled songs at learning or retrieval, then the incidence of such clustering should be equal for both scrambled and organized charts. Otherwise, conceptual clustering should be more frequent for the organized charts.

At least one other means exists by which subjects could have mentally organized the stimuli. The scrambled charts, although conceptually unorganized, did have, by definition, a spatial organization. Subjects may have chosen to organize their recall by the proximity of items to one another on the charts. "Spatial clustering" will refer to the match between the adjacency of titles in recall, and their proximity on the scrambled chart.

Each recall protocol can be scored under each of the two clustering schemes. Recall of the scrambled charts can be scored for either spatial

clustering, as noted above, or for conceptual clustering. The conceptual scheme would show to what extent unordered items are ordered in recall. Recall of the organized charts can, of course, be scored conceptually. Scores of organized charts on spatial clustering should be very low. Comparing them with spatial clustering of the scrambled charts will show to what extent purely spatial cues are used when no conceptual cues are available.

One further aspect of the clustering analysis needs to be considered. If a subject is remembering a hierarchically organized diagram in even a moderately hierarchical fashion, then different degrees of clustering should be seen for different degrees of proximity in the input. Specifically, clustering should be highest for items in the same terminal-node category, intermediate for those in the same higher level category, and lowest for those in completely different branches of the hierarchy. This pattern of clustering was expected to be strongest in the organized/conceptual combination, because of the likely dominance of conceptual over spatial links. The pattern should be less apparent in the other combinations, but to the extent it is found, will reflect hierarchical organization.

Clustering was measured by first combining subjects from both scrambled and both organized conditions. Only adjacent pairs of titles were considered. The clustering measure took into account whether a certain pair was recalled by a subject, and if so, whether the pair members were in adjacent positions. For each pair of titles, the number of subjects that had recalled both members was noted. Then, the proportion of those cases where the two songs were adjacent in recall was calculated and averaged within the appropriate category.

Each song pair was considered to be a close, medium, or far pair. These distance categories were determined by their proximity in the appropriate hierarchy. Pairs of songs in the same terminal-node category (e.g., "Hey Jude" and "Yesterday" in Figure 1) were considered as close pairs. Medium pairs had one member in the same higher order category as the other member. For example, all *Kids* songs were a medium distance away from all *Folksongs* and *Party* songs in Figure 1. All remaining pairs were far pairs. There were 141, 207, and 1,083 pairs in close, medium, and far distance categories, respectively. The distance categories were defined separately for the scrambled and organized charts. Thus, "Hey Jude" and "Row, Row, Row Your Boat" constituted a close pair under the spatial clustering scheme (Figure 2).

Results can be seen in Table 2 for conceptual and spatial clustering, measured separately for scrambled and organized diagrams. All of the chart/clustering combinations show higher clustering for close

Table 2. Clustering proportion in Experiment 2 for each pair type in each condition on Trial 1

	Clustering scheme					
	Conceptual			Spatial		
Titles	Close	Medium	Far	Close	Medium	Far
Organized	26.4	9.4	3.0	6.3	4.4	5.4
Scrambled	13.9	5.6	5.4	8.7	7.0	5.7

than for medium or far pairs; however, this effect is quite small under spatial clustering. The most pronounced clustering occurred for close pairs in the organized chart under conceptual clustering. In this condition, when two songs were recalled, on the average they occurred adjacent to one another 26.4% of the time. The only other substantial clustering is shown by close pairs in the scrambled chart under conceptual clustering. This suggests that songs with strong preexperimental associations are recalled together to some extent even when input is disorganized.

To summarize the clustering results, the organized/conceptual combination showed the most clustering of any condition. The clustering was most apparent for the close pairs, then medium, then far, suggesting a hierarchical representation of the titles. Scoring the scrambled songs along conceptual lines produced some evidence of clustering in the close pairs. This reflects an attempt to reorganize the titles into a structure based on meaning, rather than surface features.

One final note concerns the statistical significance of the results. Because the number of far pairs greatly exceeds the number of medium pairs, which exceeds the number of close pairs, by chance one would expect far pairs to occur together most frequently, followed by medium, and then close pairs. Thus the null hypothesis is exactly the reverse of what was obtained in all the conditions, and provides a fairly stringent test of the predictions of hierarchical recall.

DISCUSSION

Experiment 2 in most respects replicated the results of Experiment 1. The additional study time in Experiment 2 was apparently devoted to exploiting the mnemonic advantage conferred by the organized title scheme. Devoting extra study time to titles could account for the Title × Trial interaction if subjects spent that time learning the structure as well as the content of the title hierarchy. With extra study time, the format of the title structure (the number of branchings or

levels, for instance) could be assimilated on one trial and aid learning of the actual titles on the next trial, and so on, gaining an increasing advantage for the organized charts. It is somewhat surprising that titles, labels, and trials did not interact to produce particularly excellent recall on Trial 3 of the logical/organized chart. It appears that labels exert about the same effect wherever and whenever they occur. The limited effect of the labels may of course be due to the fact that they were generated by the experimenter after the sorting solution was found (Halpern, 1984). Although reasonable, these may not have been the only or even the best choice of labels.

The clustering analysis showed that, whenever possible, subjects relied on conceptual categories in learning and recalling even when a spatial scheme was more overtly available. Note that the conceptual clustering score of far pairs in the organized diagrams was the lowest of any distance in any diagram/scheme condition. Thus, songs far apart conceptually are recalled together even less often than scrambled songs located far apart in the diagram.

CONCLUSIONS

These studies have further supported the notion that memory for well-known tunes can be characterized as a hierarchy organized by genre. Previous studies (Halpern, 1984) tapped this knowledge in tasks of production, verification, and incidental learning. The current experiments extend the findings to an intentional learning task. When people are aware that their memory will be tested, they take advantage of the aid provided by a hierarchical arrangement of tunes. There is some evidence that subjects will attempt to create a hierarchy even when none is present, shown by the conceptual clustering in recall of the scrambled charts. When study time is limited, as in Experiment 1, subjects make good use of the category labels provided in the logical charts. With more study time, the advantage of having the titles comparable to the putative memory structure becomes more apparent.

The usefulness of having a well-organized memory for tunes is evident in many everyday situations. Besides game shows ("Name That Tune") and trivia contests, we are often asked to search our extensive musical memory. For instance, turning on the radio in the middle of some music often causes us to try to guess the identity of a tune. When we think of songs to sing with a group or for our own amusement, some form of organized access is probably used. The sheer longevity and size of our musical memory makes the existence of such an organized system plausible.

The actual content of the memory trace is one question raised by these studies. This and previous work have used tune titles as stimuli, and the assumption has been made that the titles and the actual music reside in similar places in memory. The successful use of sounded music in the verification task (Halpern, 1984, Experiment 2) lends support to this assumption. However, it may in fact be the case that tunes and their titles can be dissociated. The relationship between the access to and content of the representation of a tune remains a topic of interest in further studies of musical memory.

Notes

I would like to thank Melissa Connors, Kenneth Ramonat, and Alice Wilkinson for running these experiments, and O. Floody and L. Postman for comments on a draft of this manuscript. An earlier version of these experiments was included in a doctoral dissertation submitted to Stanford University.

Address offprint requests to Andrea R. Halpern, Psychology Department, Bucknell University, Lewisburg, PA 17837. Received for publication September 4, 1984; revision received January 22, 1985.

References

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology, 93,* 551–560.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of Genetic Psychology, 49,* 229–240.

Bower, G. H., Black, J. B., & Turner, T. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177–220.

Bower, G. H., & Clark-Meyers, G. (1980). Memory for scripts with organized vs. random presentation. *British Journal of Psychology, 71,* 368–377.

Halpern, A. R. (1984). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 496–512.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 127–172). New York: Academic Press.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika, 42,* 319–345.

Shank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9,* 77–110.